DOCUMENT RESUME

ED 212 648 TM 810 967

AUTHOR Madsen, Harold S.
TITLE Determining the Debilitative Impact of Test Anxiety.
PUB DATE [81]
NOTE 22p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *English (Second Language); Foreign Students; Higher Education; Language Proficiency; *Language Tests; *Test Anxiety; *Test Bias; *Testing Problems
IDENTIFIERS State Anxiety; Trait Anxiety

ABSTRACT

The detrimental effects of anxiety in English as a second language/foreign language (ESL/FL) are investigated. Although empirical research on the subject of ESL/FL test affect is limited, helpful insights on test anxiety exist in the psychological literature. Two constructs in the anxiety literature are considered relevant for this study: trait and state anxiety and facilitating (as well as debilitating) anxiety. Results suggest that ESI tests can be debilitating to a substantial segment of language classes and that anxiety-prone students are not evaluated as well on stressfull tests as those who are not anxiety-prone. Evidence indicates a relationship between anxiety and sex, language proficiency, and language background. Means exist to evaluate the amount of anxiety inherent in an exam for a given language group (although this information is available only after the exam has been taken). To avoid results biased in favor of students who are not anxiety-prone, it is recommended that the exam be excluded when testing comparable groups in the future. Screening tests to eliminate unnecessary anxiety together with its hidden bias is emphasized. (Author/AL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED212648

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

DETERMINING THE DEBILITATIVE IMPACT OF TEST ANXIETY

Harold S. Madsen
Brigham Young University

The purpose of this study is to assess how detrimental the effects of anxiety are in ESL language exams. Anecdotal accounts have registered preferences for or reactions against certain kinds of language tests (Savignon 1972, Groot 1976:48, Mullen 1979:188). But empirical research on the subject of ESL or FL test affect is rather limited (Jones and others 1980, Shohamy 1980, Scott and Madsen 1980).

Helpful insights on test anxiety are available however in a large body of psychological literature spanning the last two decades. One construct still being investigated is that of trait and state anxiety (Spielberger 1966, Sarason 1975, Tobias and Hedl 1972, Newmark and others 1975). Trait refers to a fairly 'stable personality characteristic,' while State is considered more transitory, tending to 'fluctuate in response to different stimuli' (Wildemuth 1977). Studies have shown among other things that girls tend to manifest higher test anxiety than boys do (Manley and Rosemeir 1972, Wilson 1973, Morris and others 1976), and that persons with low anxiety outperform those with high anxiety (Kestenbaum and Weiner 1970, Rosenzweig 1974, Ohlenkamp 1976).

TM 810967

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. S. Madsen
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Another construct in the anxiety literature has particular relevance for this study. It had generally been assumed that anxiety was harmful or debilitating to test performance. But two decades ago Alpert and Haber identified facilitating as well as debilitating anxiety, and they produced an instrument to identify the two types (1960). Subsequent studies have broadened our understanding of facilitating and debilitating anxiety, relating them for example to students' general outlook on life and performance in school (Gaudry and Spielberger 1971, Bronzaft and Dobrow 1976, Bronzaft and others 1973, Scovel 1978).

The empirical research on second-language test anxiety referred to at the outset has demonstrated expected differences between students at various levels of proficiency and has shed light on the result of practice effect, credit versus non-credit contexts, and even the anxiety profiles of A, B, C, D, and E students. A consistent finding has been the dramatic contrasts between the amount of State anxiety generated by various types of language tests (Stevenson 1979, Shohamy 1980, Madsen and others 1981). Implicit in these findings is the suggestion that language exams which generate considerably more anxiety than others be avoided. A contrasting view, however, is that such a step would be premature until we determine whether or not these tests are in fact debilitating rather than facilitating. Oller (1980) has suggested, for instance, that if the cloze test were as debilitating as some suggest, it would not produce the robust correlations it does with other measures.

The object of this investigation is to discover some means of assessing when the anxiety level of a test crosses the debilitating threshold. Since the intercorrelation of well-written language exams is generally rather high, it seems plausible to assume that many students are not excessively disoriented by higher-anxiety-producing exams. If any of these tests are in fact debilitating, they probably disturb the performance of only a portion of the student population. Identifying this population should enable the researcher to determine to what extent a language exam is biased against that group. There is some evidence that persons with certain cultural traits may be more susceptible to anxiety on a given language test than others are (Barabasz 1970, Bronzaft and others 1974, Scott and Madsen 1980). And we have already mentioned the relationship between sex and anxiety as well as proficiency level and anxiety.

In looking for that group which might be particularly susceptible to anxiety-arousing tests, one is led by the psychological literature to study people with high Trait anxiety: that is, people who are by nature prone to be anxious when taking any kind of examination. It is hypothesized that students who are anxiety prone will not be evaluated as well on stressful tests as are those who are not anxiety prone. One rationale for this assumption is that anxiety-prone individuals engage in more task-irrelevant activities while taking a test (Wine 1971, Nottelmann 1975, Dusek and others 1976). It is assumed that if this tendency

is heightened during highly frustrating tests, there will be more disorientation and irrelevant behavior, resulting in a less accurate measure of language ability. Less accurate measurement should in turn be reflected in a depressed correlation with a general (non-threatening) measure of language proficiency, particularly in contrast to the performance of a low-anxiety group. Also, the literature suggests that the females in this study and those with less proficiency would register higher levels of anxiety.

Method

Subjects. The 146 adult subjects involved in this study are enrolled at five levels of instruction in the Brigham Young University English Language Center intensive English program. Of these, 45 males and 69 females (a total of 114) completed all of the testing. This group consists of 74 Spanish speakers, 25 Japanese speakers, and 15 from other language groups (6 French, 2 Chinese, and one each of Portuguese, Finish, Greek, German, Arabic, Serbo-Croatian, and Indonesian). They range in proficiency from near beginning level to advanced (530 on the TOEFL).

Measurement. Six weeks prior to the end of the term, all students were administered the BYU/ELC Progress Battery. This five-part instrument included a twenty-item oral interview conducted individually, a thirty-item multiple-choice grammar

test, a ten-item true-false reading comprehension test based on two prose selections, a single dictation, and a twenty-item multiple-choice, appropriate-paraphrase listening test.

As a measure of their Trait anxiety, all students were administered the Alpert-Haber Achievement Anxiety Test, the last week of the semester. This test was translated into Spanish and Japanese. The fifteen speakers of other languages took the English version, with a teacher available to answer questions. The AAT includes 10 items that measure debilitating anxiety (such as 'Nervousness while taking an exam or test hinders me from doing well'); 10 items that measure facilitating anxiety (for example, 'I work most effectively under pressure, as when the task is very important'); plus neutral items; and all three types are scrambled. Responses are marked on a five-point Likert-type scale, ranging from 'almost never' to 'almost always.'

At the end of the semester students were administered the six-section BYU/ELC Promotion Battery. This included a new twenty-item oral interview conducted individually, a thirty-item multiple-choice grammar test, a thirty-item multiple-choice sentence-paraphrase reading test, a single dictation, ten multiple-choice appropriate paraphrase listening comprehension items, and two selected-deletion cloze passages totalling thirty blanks.

Immediately upon completing the Promotion Battery, students were administered a short three-item State Anxiety questionnaire related to each of the six subtests they had just taken. For

each test they would respond on a five-point Likert-type scale (ranging from 'strongly agree' to 'strongly disagree') on the following statements: 1) I liked the test, 2) This test was difficult, 3) I felt pleasant (happy, calm) during this test. These are adaptations of the items that factor analysis had shown to be highly emotive on the Jones-Madsen Affect Questionnaire (Jones and others 1980).

Data analysis. The average State anxiety level was determined for each of the six parts of the Promotion Battery. High and low anxiety prone students were selected with reference to self ratings on the AAT (the top third being rated as high Trait anxiety subjects, the low third being rated as low Trait anxiety subjects). Each subtest of the Promotion battery was correlated with the Promotion Test total, and a test of difference between two Pearson coefficients from independent samples was calculated, including a correction for attenuation. The effects of sex, language background, and proficiency level were evaluated for the Trait and State anxiety measures, utilizing analysis of covariance.

Results and Discussion

State anxiety ratings of the six subtests (Table 1) rather consistently identified the oral interview as the least anxiety

producing and the reading test as most anxiety producing. On an absolute scale, the reading test was the only measure that consistently registered in the +Anxiety range (above 9.0). Cloze was next most frustrating. Smallest differences between tests of high and low frustration are recorded at levels one and two. This may be due in part to the limited ability of students at these levels to interpret the English-language questions (simple as they were).

As predicted, means on the Promotion Battery for the low anxiety group were uniformly higher than for the high anxiety group, all differences but the oral interview and the writing test being statistically significant (Table 2). Reliabilities for the subtests were computed using Kuder-Richardson formula 21 (Table 3).[1] Reliability coefficients for the writing test (dictation) and the total utilized the parallel test form procedure (Thorndike 1971:404-406), i.e. Pearson correlations between Promotion and Progress Battery dictations, and between Promotion and Progress totals, minus the cloze.

For high and low anxiety students, Pearson product-moment correlations were calculated for each of the six Promotion subtests with the Progress Battery total (Table 4). These were also corrected for attenuation. A statistically significant difference (at the .027 level) was found between the high and low anxiety group correlations on the frustrating reading test. This supports the principal hypothesis that a significantly

lower correlation will result for high Trait anxiety students on tests that maximize State anxiety.

While one might have expected to find differences among one or more other tests that tend to generate State anxiety, there is a rather straight forward explanation as to why only the reading test produced this result: It was the only subtest that crossed into the + Anxiety threshold on an absolute scale. The cloze and listening tests produced more anxiety than did speaking, grammar, and writing; but they registered only in the twilight zone between anxiety producing and anxiety allaying measures.

An analysis of covariance was next employed (Table 5) to evaluate the effect of proficiency level, sex, and language on Trait anxiety--the general tendency toward anxiety as measured by the AAT. Proficiency (as measured on a combined Progress-Promotion Battery total) proved to be the only significant variable. However, a chi square test evaluating the effect of sex on high and low Trait anxiety ratings revealed a statistically significant difference (chi square = 3.86), females being more inclined toward the high anxiety rating.

Analysis of covariance was also used (Table 6) to evaluate the effect of sex, language and proficiency on the State anxiety measures. A total State anxiety measure was generated for each individual by combining the anxiety ratings he or she gave to each of the six subtests. Results found both sex and language to be significant, with proficiency approaching significance. Females

indicated that they experienced greater State anxiety than the males said they did; Japanese registered higher State anxiety than did Spanish speakers. These latter findings tend to corroborate what we find in the literature.

Conclusions

Results of this study suggest that an ESL test can be debilitating to a substantial segment of our classes. While the frustrating test in this study was a reading exam, we obviously cannot generalize these results to all reading measures or even to all reading tests of this type. It is apparent, however, that relatively simple means are at our disposal to evaluate the amount of frustration or anxiety inherent in an exam for a given language group (although this information is available to us only after the exam has been taken). To avoid results that are biased in favor of students who are not anxiety prone, that exam could be excluded when testing comparable groups in the future.

This investigation has attempted to determine the degree of frustration encountered on a variety of ESL language exams, but not the cause of this frustration. Certainly the source of frustration is of the utmost importance, and this also needs to be determined. Anecdotal accounts and research indicate that in addition to the form of the exam (cloze, oral interview, etc.) faulty instructions, lack of face validity, difficulty level,

insufficient time, and flagrant cheating by other students are a few of the factors that can cause anxiety and frustration while taking a test.

In further investigations of debilitating test anxiety, it would be helpful to consider other populations as well, such as national or ethnic groups, those with differing learning styles, etc. Also there is no need to restrict such studies to classical empirical research. For example, personal interviews following the exam may provide excellent insights not otherwise available.

Finally, this study indicates the value in looking beyond the traditional test criteria of validity and reliability. There is clearly a need to screen our tests to eliminate unnecessary anxiety together with its hidden bias.

Level	Category					
	writing	oral	grammar	reading	listening	cloze
1	7.0	6.9	7.5	8.0	7.6	9.2
2	8.1	7.6	8.1	9.2	8.7	8.4
3	7.8	7.7	7.1	10.7	8.4	9.6
4	8.9	7.6	8.7	11.1	9.5	8.6
5	7.6	5.9	6.6	10.5	8.4	8.2
Average	7.88	7.14	7.60	9.90	8.52	8.80

Note: Absolute scale 3 = anxiety allaying
9 = anxiety threshold
15 = highly anxiety producing

Table 1

State Anxiety Ratings by Level for the Six Subtests

Anxiety Group	N	Category					
		writing	oral	grammar	reading	listening	cloze
High Anxiety	42	76.4	62.4	47.6	28.5	55.0	42.1
Low Anxiety	38	78.2	66.9	59.8 *	42.7 *	64.2 *	51.3 *

* Between-group difference: $P = < .05$

Table 2

Means of High and Low Trait Anxiety Groups On the Six Subtests

Anxiety Group	N	Category					
		writing	oral	grammar	reading	listening	cloze
High Anxiety	42	.720	.690	.892	.657	.700	.780
Low Anxiety	38	.700	.820	.923	.892	.570	.810

* Spearman-Brown correction for uniform item no. of 30

Table 3

Reliabilities (KR-21) of High and Low Trait Anxiety Groups On the Six Subtests*

Anxiety Group	N	Category					
		writing	oral	grammar	reading	listening	cloze
High Anxiety	42	.62	.68	.80	.57	.64	.71
Low Anxiety	38	.43	.65	.85	.82 *	.60	.71

* Difference between correlations: P = $<$.05 (.027)

(Note: When correlations are corrected for attenuation, the only significant difference is still reading [.03] .)

Table 4

Pearson Correlations of High and Low Trait Anxiety Groups On the Six Subtests

Source	DF	SS	F	P
Proficiency	1	286.8477	6.18	.0145
Sex	1	84.4638	1.82	.1802
Language	2	57.0851	0.61	.5426
Sex x Language	2	25.9050	0.28	.7571
Error	107			

Table 5

Analysis of Covariance

Dependent Variable: Trait Anxiety

Source	DF	SS	F	P
Sex	1	360.0792	4.46	.0370
Language	2	1438.9434	8.92	.0003
Sex x Language	2	103.4662	0.64	.5287
Proficiency	1	260.3537	3.23	.0753
Proficiency x Sex	1	246.5128	3.06	.0834
Proficiency x Language	2	638.9399	3.96	.0220
Error	104			

Table 6

Analysis of Covariance

Dependent Variable: State Anxiety

Footnotes

[1]KR-21 assumes not only that the tests are unspeeded but also that items are equally difficult (a questionable assumption on most language tests). To the extent that items vary significantly in difficulty, KR-21 provides a lower reliability estimate than does KR-20 (Thorndike 1971:413-415).

References

Alpert, R. and R. N. Haber. 1960. "Anxiety in academic achievement situations." Journal of Abnormal Psychology 61.207-215.

Barabasz, Arreed F. 1970. "Galvanic skin response and test anxiety among negroes and caucasians." Child Study Journal 1.33-35.

Bronzaft, Arline L. and Stephan B. Dobrow. 1976. "Test anxiety, and the economic recession." Psychological Reports 38.1211-1215.

Bronzaft, Arline L. and others. 1973. "Test anxiety and cheating on college examinations." Psychological Reports 32.149-150.

Bronzaft, Arline L. and others. 1974. "Test anxiety among black college students: a cross-cultural study." Journal of Negro Education 43.190-193.

Dusek, Jerome B. and others. 1976. "Attention, encoding, and information processing in low- and high-test-anxious children." Child Development 47.201-207.

Gaudry, Eric and Charles D. Spielberger (eds.). 1971. Anxiety and educational achievement. New York: Wiley.

Groot, P. J. M. 1976. Luistervaardigheid, Frans, Duits, Engels. Doelstelling en toetsing. Amsterdam: Meulenhoff.

Jones, Randall L. and others. 1980. "Evaluating affective variables in second-language testing." Paper presented at the Fourteenth Annual TESOL Convention, San Francisco, California.

Kestenbaum, Joel M. and Bernard Weiner. 1970. "Achievement performance related to achievement motivation and test anxiety." Journal of Consulting and Clinical Psychology 34.343-344.

Madsen, Harold S. and others. 1981. "Evaluating student attitudes toward second language tests." (forthcoming)

Manley, Merlin J. and Robert A. Rosemeir. 1972. "Developmental trends in general and test anxiety among junior and senior high school students." Journal of Genetic Psychology 120.219-226.

Morris, Larry W. and others. 1976. "Components of school anxiety: Developmental trends and sex differences." Journal of Genetic Psychology 128.49-57.

Mullen, Karen A. 1979. "An alternative to the cloze test." On TESOL '79. Yorio, Carlos A. and others, eds. Washington, D.C.: TESOL.

Newmark, Charles S. and others. 1975. "Test-induced anxiety with children." Journal of Personality Assessment 39.400-413.

Nottelmann, Edith D. 1975. "Test anxiety and off-task behavior in evaluation situations." ED 113 022.

Ohlenkamp, Edward A. 1976. "The relationship among test anxiety, self-esteem, and achievement for cooperative career education students in grades eleven and twelve." Ed.D. dissertation, Northern Illinois Univ.

Oller, John W., Jr. 1980. Private communication.

Rosenzweig, Steven Jeffrey. 1974. "The effects of examiner anxiety level, student anxiety level, and examiner-student sex interaction, on student performance in a group test-taking situation." Ed.D dissertation, Boston University School of Education.

Sarason, Irwin C. 1975. "Test anxiety and the self-disclosing coping model." Journal of Consulting and Clinical Psychology 43.148-153.

Savignon, Sandra J. 1972. Communicative competence: an experiment in foreign language teaching. Philadelphia: The Center for Curriculum Development, Inc.

Scott, Mary Lee and Harold S. Madsen. 1980. "The influence of retesting on test affect." Paper presented at the 1980 Testing Conference, TESOL Summer Institute; Albuquerque, New Mexico.

Scovel, Thomas. 1978. "The effect of affect on foreign language learning: a review of the anxiety research." Language Learning 28.99-104.

Shohamy, Elana. 1980. "Students' attitudes toward tests: affective considerations in testing." Paper presented at the Fourteenth Annual TESOL Convention, San Francisco, California.

Spielberger, Charles D. 1966. "Theory and research on anxiety." In Spielberger, Charles D. (ed.) Anxiety and Behavior. New York: Academic Press.

Stevenson, Diana G. 1979. "The experimental evaluation of test affect." M.A. thesis, Brigham Young University.

Thorndike, Robert L. (ed.). 1971. Educational measurement. Second Edition. Washington, D.C.: American Council on Education.

Tobias, Sigmund and John J. Hedl, Jr. 1972. "Test anxiety: situationally specific or general?" ED 058 853.

Wildemuth, Barbara. 1977. (Introduction) "Test anxiety: an extensive bibliography." TM Report 64. ED 152 860.

Wilson, Shirley Byrd. 1973. "Self-esteem and test anxiety in black inner-city junior high youngsters." Ph.D. dissertation, University of Pittsburgh.

Wine, Jeri. 1971. "Test anxiety and direction of attention." Psychological Bulletin 76.92-104.